Monitoring Multinomial Logit Profiles Via Log-Linear Models

R. Noorossana*, A. Saghaei, H. Izadbakhsh & O. Aghababaei

Rassoul Noorossana, Industrial Engineering Department, Iran University of Science and Technology, rassoul@iust.ac.ir
Abbas Saghaei, Industrial Engineering Department, Islamic Azad University, Science and Research Branch, a.saghaei@srbiau.ac.ir
Hamid Izadbakhsh, Industrial Engineering Department, Iran University of Science and Technology Tehran, Iran, hizadbakhsh@iust.ac.ir
Omid Aghababaei, Statistics Department, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

KEYWORDS
Loglinear Models, Average Run Length (ARL), Multivariate Exponentially Weighted Moving Average (MEWMA), Multinomial Logit Regression, Profile Monitoring

ABSTRACT
In certain statistical process control applications, quality of a process or product can be characterized by a function commonly referred to as profile. Some of the potential applications of profile monitoring are cases where quality characteristic of interest is modelled using binary, multinomial or ordinal variables. In this paper, profiles with multinomial response are studied. For this purpose, multinomial log it regression (MLR) is considered as the basis. Then, the MLR is converted to Poisson GLM with log link. Two methods including Multivariate exponentially weighted moving average (MEWMA) statistics, and Likelihood ratio test (LRT) statistics are proposed to monitor MLR profiles in phase II. Performances of these three methods are evaluated by average run length criterion (ARL). A case study from alloy fasteners manufacturing process is used to illustrate the implementation of the proposed approach. Results indicate satisfactory performance for the proposed method.


1. Introduction
In the last 2 decades, the control charts for cases, in which the quality of a process or product is characterized by a function between two or more variables, are developing both in theory and application. So far, several investigations have been studied on various types of this function, which called “profile”, including linear and nonlinear profiles. In addition, profile monitoring has been applied in some real applications such as automotive [1], piece manufacturing [2, 3], Calibration [4], Tourism [5], Economy [6], and Charging and loading process [7]. Many authors including [8-16] worked on linear profiles. Multiple and polynomial linear profile are considered by [1, 17-21]. Ref. [22-27] studied nonlinear profile models. Multivariate linear profiles are investigated by [4, 7, 28, 29]. For more details, see Noorossana et al. [30]. The continuity of the response variables is assumed in all aforementioned researches. However, in many applications the response variables could be binary, multinomial or ordinal. These situations are the most well-known especial cases of general linear models(GLMs) family and often model with logistic regression. Binary logistic profile was studied by [2, 31] Phase I and Phase II, respectively. Noorossana and Izadbakhsh [3] investigated multinomial logistic profile in Phase I and this profile in Phase II was investigated by Noorossana et al. [32]. Izadbakhsh et al. in [33] studied ordinal logistic profile in Phase II for tourism case studies. Noorossana and Izadbakhsh discussed [34] ordinal logistic profile in Phase I.

* Corresponding author: Rassoul Noorossana
Email: rassoul@iust.ac.ir
Our focus in this paper is to study multinomial logistic profiles and to propose two methods in Phase II including MEWMA, and LRT. Using Poisson GLM with log link instead of multinomial logistic is the novelty of this paper. Hence, estimating the coefficients became easier and more accurate. Moreover, monitoring the strength of alloy fastener as a case study is presented. To evaluate the performances of the proposed methods, average run length (ARL) measure is used.

The reminder of this paper is organized as follows. In Section 2, the multinomial logistic regression model and the methods used to estimate its parameters are discussed. MEWMA and LRT methods are presented in Section 3. In Section 4, the performance of the two profile monitoring approaches are evaluated based on data from the case study. Our concluding remarks are provided in the Section 4.

2. Multinomial Logistic Regression

There are two types of response variables, continues and discrete. The discrete variables divided in to counts variables and categorical variables. Multinomial, ordinal and binary are the three sub categories of categorical variables. This concept is demonstrated in Figure 1. Multinomial response variable is considered in this paper. The major difference between multinomial and ordinal is that the multinomial variables do not have any natural order. In other words, if response variable has more than two levels it may have ordered nature, such as weak, moderate, and good, or it may have unordered nature, such as product categories.

Consider a multinomial response variable with $k$ - categorical levels and a $p$ - dimensional vector of covariate variables denoted by $\mathbf{X}$. For multinomial logistic regression modelling, it is very common to consider the $1^{st}$ level as the basic level and to calculate other levels based on this base. In other words, multinomial logistic regression model could be defined by [36]:

$$\pi_{ij} = \exp \left( \beta_{j}' \mathbf{x}_i \right), i = 1,..., n, j = 1,..., k$$

(1)

where, $n$ is the number of covariate’s settings, $\beta_j = 0$ and $\pi_{ij} = \Pr(Y_{ij} = j) = \Pr(Y_{ij}) \forall i = 1,..., n, j = 1,..., k$.

These equations are known as logit links and since, $\pi_{i1} = 1 - \sum_{j=2}^{k} \pi_{ij}$, we would have $k - 1$ logit equations.

There are other links such as probit or log-log in literature, however the logit link is more common. Moreover, It could be shown easily that:

$$\pi_{ij} = \frac{\exp(\beta_j' \mathbf{x}_i)}{\sum_{j=1}^{k} \exp(\beta_j' \mathbf{x}_i)}, j = 1,..., k$$

(2)

where, $\eta_{ij} = \beta_j' \mathbf{x}_i$.

Let consider $(y_{11t}, ..., y_{kt})$ be the $t^{th}$ profile sample that are shown in the following form:

<table>
<thead>
<tr>
<th>Explanatory Variables (i)</th>
<th>Response Category (j)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$y_{11t}$</td>
<td>$y_{1kt}$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$y_{21t}$</td>
<td>$y_{2kt}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$x_n$</td>
<td>$y_{n1t}$</td>
<td>$y_{nkt}$</td>
</tr>
</tbody>
</table>

Consider $y_{ijt}$ as independently Poisson variable with mean $\lambda_{ijt}$. As $\sum_{j=1}^{k} y_{ijt}$ is Poisson with mean $\sum_{j=1}^{k} \lambda_{ijt}$ it could be shown easily that the conditional distribution of the $y_{ijt}$ given total counts $\sum_{j=1}^{k} y_{ij}$ is multinomial.

In other words, the likelihood associated with multinomial observations is the same as the likelihood associated with Poisson GLM with log link which are constrained by a fixed total.

Formally, if we define $\lambda_{ijt} = \lambda_{ij} \exp \left( \beta_{j}' \mathbf{x}_i \right)$, the log-linear model could be as:

$$\log \lambda_{ijt} = \log \lambda_{ij} + \beta_{j}' \mathbf{x}_i$$

$$= \log(\phi_i) + \beta_{j}' \mathbf{x}_i$$

(3)
Estimating the parameters \((\phi, \beta)\) is done using a two-stage log likelihood maximization with the following steps (Algorithm 1).

**Step 1:** Consider log likelihood as:

\[
l(\phi, \beta) = \sum_{i=1}^{n} \left( -\alpha_i + y_i \log \lambda_{ij} \right)
\]

(4).

**Step 2:** Maximize Equation (4) based on \(\phi\).

So,

\[
\frac{\partial l}{\partial \phi} = \sum_{j=1}^{k} \left( -\exp(\beta_j'x_i) + \frac{y_{ij}}{\phi} \right)
\]

(5)

and \(\hat{\phi}\) could be estimated using:

\[
\hat{\phi} = \frac{\sum_{j=1}^{k} y_{ij}}{\sum_{j=1}^{k} \exp(\beta_j'x_i)}
\]

(6)

**Step 3:** Maximize Equation (4) based on \(\beta_j\).

To do so, consider \(\phi\) as fixed and to converge more rapidly, without loss of generality, consider the log likelihood for each category level as:

\[
l_j(\beta_j) = \sum_{i=1}^{n} \left( -\alpha_i + y_i \log \lambda_{ij} + \frac{y_{ij}}{\phi} \right)
\]

(7)

So, the first and the second derivative of Equation (7) are:

\[
\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \left( -\phi \alpha_i \exp(\beta_j'x_i) + \frac{y_{ij} x_i}{\phi} \right)
\]

(10)

\[
\frac{\partial^2 l}{\partial \beta_j^2} = \sum_{i=1}^{n} \left( -\phi x_i^2 \exp(\beta_j'x_i) \right)
\]

(11)

**Step 4:** Calculate the new estimate of \(\beta_j\) based on Newton-Raphson method.

**Step 5:** After fitting models to each category in step 4, new estimate of \(\phi\) is calculate using step 2.

3. Multinomial Logistic Profiles Monitoring

**MEWMA Approach**

In order to monitor \(\xi = (\phi, \beta)\), we use MEWMA control chart which is proposed by [34]:

\[
Z_{t, \xi} = \theta(\hat{\xi} - \bar{\xi}) + (1 - \theta)Z_{(t-1), \xi}
\]

(8)

where \(\theta(0 < \theta < 1)\) is a smoothing parameter. When the process is in-control \(Z_{t, \xi}\) has mean vector zero and covariance matrix \(\Sigma_{Z_{t, \xi}} = \left( \frac{\theta}{2 - \theta} \right) \Sigma_{\xi} \). Based on these assumptions, \(T^{2}_{z, \xi}\), the monitoring statistic, is given by:

\[
T^{2}_{z, \xi} = Z_{t, \xi} \Sigma_{t, \xi}^{-1} Z_{t, \xi}'
\]

(9)

Obviously, the process is in-control when \(T^{2}_{z, \xi} \leq h_M\).

\(h_M\) is obtained by using a Monte Carlo simulation.

4. The LRT Approach

According to Agresti (2003), likelihood ratio test (LRT) is a common test to evaluate goodness of fit test in generalized linear models (GLMs) literature. In the LRT approach, the ratio of maximum likelihoods based on the reference and estimated models are compared. If \(\tilde{\xi} = (\hat{\phi}, \hat{\beta})\) denotes the maximum likelihood estimates which is obtained by Algorithm 1 for the model parameters vector, \(\xi = (\phi, \beta)\), then the two log-likelihood functions can be defined as:

\[
l(\phi, \beta; y_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{k} (-\lambda_{ij} + y_{ij} \log \lambda_{ij})
\]

(10)

\[
l(\hat{\phi}, \hat{\beta}; y_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{k} (-\hat{\lambda}_{ij} + y_{ij} \log \hat{\lambda}_{ij})
\]

(11)

By using equations (10) and (11), the proposed control chart uses the following statistic:

\[
\lambda_t = 2 \left( \log L(\tilde{\xi}; y_i) - \log L(\xi; y_i) \right)
\]

(12)
This statistic, which is usually referred to as deviance, follows an approximate chi-square distribution with \( r \) degrees of freedom. \( \lambda \) is equal to the difference in the degrees of freedom of the reference model and the estimated model. Therefore, the recommended upper control limit for this chart would be \( \chi^2_{2r} \). It is obvious that large values of \( \lambda \) indicate the out-of-control conditions.

5. The Case Study

The compressive strength of an alloy fastener that used in many applications such as aircraft construction is a very important quality characteristic. For controlling this characteristic, we have to control the strength for different loadings as measured in pounds per square inch (psi). This is the certain type of profile. The state of alloy fasteners when testing with certain load is denoted by response variable. In this case, the response variable has three levels, major defect, minor defect, and no defect. The covariate \( (X) \) is the loading strength. Based on historical data, the model is defined as follows:

\[
\begin{align*}
\log \left( \frac{\pi_1}{\pi_3} \right) &= 2 + \log (x_i) \\
\log \left( \frac{\pi_2}{\pi_3} \right) &= 1.5 + 2 \log (x_i)
\end{align*}
\]

where, \( \pi_1, \pi_2, \text{and} \, \pi_3 \) are the probability of having major, minor, or no defect in certain loads in cluding \( x_i = \{2500, 2700, \ldots, 4300\} \).

The two approaches are simulated for 10,000 times and the results are compared by using average run length (ARL). The in-control ARL \((ARL_0)\) of 200 is very common among different statistical process control schemes. Hence, the control limits are set to gain \( ARL_0 = 200 \). We, then evaluate the performance of the approaches by considering shifts in the parameters as a multiple of the parameter standard deviation such as \( \beta_{s1} + \lambda \sigma_{\beta_{s1}} \). Table 1 and Table 2 show the ARL values for different shifts in the parameters of the proposed approaches for \( n=30 \) and \( n=100 \).

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{s1} ) shifts to</td>
<td>MEWMA</td>
<td>198.0</td>
<td>24.03</td>
<td>7.36</td>
<td>4.27</td>
<td>3.08</td>
<td>2.44</td>
<td>2.09</td>
<td>1.87</td>
<td>1.70</td>
</tr>
<tr>
<td>( \beta_{s1} + \lambda \sigma_{\beta_{s1}} )</td>
<td>LRT</td>
<td>199.5</td>
<td>101.41</td>
<td>28.63</td>
<td>9.27</td>
<td>3.76</td>
<td>1.98</td>
<td>1.34</td>
<td>1.11</td>
<td>1.03</td>
</tr>
<tr>
<td>( \beta_{s3} ) shifts to</td>
<td>MEWMA</td>
<td>198.0</td>
<td>55.85</td>
<td>10.84</td>
<td>5.31</td>
<td>2.78</td>
<td>1.93</td>
<td>1.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{s1} + \lambda \sigma_{\beta_{s1}} )</td>
<td>LRT</td>
<td>199.5</td>
<td>105.18</td>
<td>30.80</td>
<td>8.89</td>
<td>3.45</td>
<td>1.84</td>
<td>1.24</td>
<td>1.05</td>
<td>1.00</td>
</tr>
<tr>
<td>( \beta_{s2} ) shifts to</td>
<td>MEWMA</td>
<td>198.0</td>
<td>57.60</td>
<td>12.15</td>
<td>6.13</td>
<td>4.13</td>
<td>2.56</td>
<td>2.31</td>
<td>1.99</td>
<td>1.86</td>
</tr>
<tr>
<td>( \beta_{s2} + \lambda \sigma_{\beta_{s2}} )</td>
<td>LRT</td>
<td>199.5</td>
<td>121.10</td>
<td>43.48</td>
<td>15.51</td>
<td>6.60</td>
<td>3.25</td>
<td>1.97</td>
<td>1.43</td>
<td>1.17</td>
</tr>
<tr>
<td>( \beta_{s3} ) shifts to</td>
<td>MEWMA</td>
<td>198.0</td>
<td>29.68</td>
<td>10.38</td>
<td>5.98</td>
<td>4.19</td>
<td>3.30</td>
<td>2.71</td>
<td>2.34</td>
<td>2.09</td>
</tr>
<tr>
<td>( \beta_{s2} + \lambda \sigma_{\beta_{s2}} )</td>
<td>LRT</td>
<td>199.5</td>
<td>136.98</td>
<td>66.90</td>
<td>32.82</td>
<td>16.19</td>
<td>8.89</td>
<td>5.32</td>
<td>3.63</td>
<td>2.63</td>
</tr>
</tbody>
</table>

| \( \beta_{s2} \) shifts to | MEWMA | 198.0 | 39.12 | 9.90 | 5.29 | 3.63 | 2.83 | 2.34 | 2.07 | 1.90 | 1.75 |
| \( \beta_{s2} + \lambda \sigma_{\beta_{s2}} \) | LRT | 201.0 | 118.96 | 41.30 | 14.30 | 5.80 | 2.89 | 1.75 | 1.30 | 1.12 | 1.03 |
| \( \beta_{s3} \) shifts to | MEWMA | 198.0 | 30.64 | 10.07 | 5.68 | 3.99 | 3.11 | 2.59 | 2.24 | 2.01 | 1.83 |
| \( \beta_{s3} + \lambda \sigma_{\beta_{s3}} \) | LRT | 201.0 | 162.38 | 95.71 | 25.54 | 13.88 | 8.04 | 4.92 | 3.26 | 2.32 | 1.79 |

International Journal of Industrial Engineering & Production Research, June 2013, Vol. 24, No. 2
6. Conclusions
In this paper, two approaches were proposed to monitor multinomial logistic profiles in Phase II. Multinomial responses are commonly encountered in situations where data can be classified as categories. In our study, we used the case of alloy fasteners to evaluate performances of the two proposed approaches using simulation. Performance of the proposed method are evaluated in terms of average run length criterion. Novelty of this paper is in using Poisson GLM with log link.

References

[26] Vaghefi, A., Tajbakhsh, S., Noorossana, R., Phase II Monitoring of Nonlinear Profiles. Communications in


